Rezarta Islamaj DoÇ§an

List of Publications by Year
in descending order

35 papers	2,144 citations	430874 18 h-index	454955 30 g-index
35 all docs	35 docs citations	35 times ranked	2449 citing authors

#	Article	IF	Citations
1	NCBI disease corpus: A resource for disease name recognition and concept normalization. Journal of Biomedical Informatics, 2014, 47, 1-10.	4.3	525
2	DNorm: disease name normalization with pairwise learning to rank. Bioinformatics, 2013, 29, 2909-2917.	4.1	436
3	SplicePort--An interactive splice-site analysis tool. Nucleic Acids Research, 2007, 35, W285-W291.	14.5	190
4	Understanding PubMed(R) user search behavior through log analysis. Database: the Journal of Biological Databases and Curation, 2009, 2009, bap018-bap018.	3.0	189
5	BioC: a minimalist approach to interoperability for biomedical text processing. Database: the Journal of Biological Databases and Curation, 2013, 2013, bat064-bat064.	3.0	123
6	The Protein-Protein Interaction tasks of BioCreative III: classification/ranking of articles and linking bio-ontology concepts to full text. BMC Bioinformatics, 2011, 12, S3.	2.6	121
7	Semi-automatic semantic annotation of PubMed queries: A study on quality, efficiency, satisfaction. Journal of Biomedical Informatics, 2011, 44, 310-318.	4.3	68
8	Author name disambiguation for <scp>P</scp>ub<scp>M</scp>ed. Journal of the Association for Information Science and Technology, 2014, 65, 765-781.	2.9	68
9	PMC text mining subset in BioC: about three million full-text articles and growing. Bioinformatics, 2019, 35, 3533-3535.	4.1	51
10	TeamTat: a collaborative text annotation tool. Nucleic Acids Research, 2020, 48, W5-W11.	14.5	34
11	Overview of the BioCreative VI Precision Medicine Track: mining protein interactions and mutations for precision medicine. Database: the Journal of Biological Databases and Curation, 2019, 2019, .	3.0	30
12	An Evolutionary Algorithm Approach for Feature Generation from Sequence Data and Its Application to DNA Splice Site Prediction. IEEE/ACM Transactions on Computational Biology and Bioinformatics, 2012, 9, 1387-1398.	3.0	29
13	BioCreative V BioC track overview: collaborative biocurator assistant task for BioGRID. Database: the Journal of Biological Databases and Curation, 2016, 2016, baw121.	3.0	28
14	Evolving use of ancestry, ethnicity, and race in genetics research—A survey spanning seven decades. American Journal of Human Genetics, 2021, 108, 2215-2223.	6.2	27
15	Extracting Rx information from clinical narrative. Journal of the American Medical Informatics Association: JAMIA, 2010, 17, 536-539.	4.4	26
16	NLM-Chem, a new resource for chemical entity recognition in PubMed full text literature. Scientific Data, 2021, 8, 91.	5.3	26
17	The BioC-BioGRID corpus: full text articles annotated for curation of protein—protein and genetic interactions. Database: the Journal of Biological Databases and Curation, 2017, 2017, baw147.	3.0	24
18	Features generated for computational splice-site prediction correspond to functional elements. BMC Bioinformatics, 2007, 8, 410.	2.6	21

#	Article	IF	Citations
19	BioC interoperability track overview. Database: the Journal of Biological Databases and Curation, 2014, 2014, bau053-bau053.	3.0	15
20	A context-blocks model for identifying clinical relationships in patient records. BMC Bioinformatics, 2011, 12, S3.	2.6	14
21	Structural footprinting in protein structure comparison: the impact of structural fragments. BMC Structural Biology, 2007, 7, 53.	2.3	13
22	Finding abbreviations in biomedical literature: three BioC-compatible modules and four BioC-formatted corpora. Database: the Journal of Biological Databases and Curation, 2014, 2014, bau044-bau044.	3.0	12
23	Click-words: learning to predict document keywords from a user perspective. Bioinformatics, 2010, 26, 2767-2775.	4.1	11
24	Crowdsourcing and curation: perspectives from biology and natural language processing. Database: the Journal of Biological Databases and Curation, 2016, 2016, baw115.	3.0	10
25	PubMed Text Similarity Model and its application to curation efforts in the Conserved Domain Database. Database: the Journal of Biological Databases and Curation, 2019, 2019, .	3.0	10
26	NLM-Gene, a richly annotated gold standard dataset for gene entities that addresses ambiguity and multi-species gene recognition. Journal of Biomedical Informatics, 2021, 118, 103779.	4.3	9
27	A Feature Generation Algorithm for Sequences with Application to Splice-Site Prediction. Lecture Notes in Computer Science, 2006, , 553-560.	1.3	8
28	Natural language processing pipelines to annotate BioC collections with an application to the NCBI disease corpus. Database: the Journal of Biological Databases and Curation, 2014, 2014, bau056-bau056.	3.0	8
29	BioC implementations in Go, Perl, Python and Ruby. Database: the Journal of Biological Databases and Curation, 2014, 2014, bau059-bau059.	3.0	8
30	BioCreative VI Precision Medicine Track: creating a training corpus for mining protein-protein interactions affected by mutations. , 2017, , .		6
31	Topics in machine learning for biomedical literature analysis and text retrieval. Journal of Biomedical Semantics, 2012, 3, S1.	1.6	2
32	Characterizing RNA Secondary-Structure Features and Their Effects on Splice-Site Prediction. , 2007, , .		1
33	A Textual Representation Scheme for Identifying Clinical Relationships in Patient Records. , 2010, 2010, 995-998.		1
34	Topics in machine learning for biomedical literature analysis and text retrieval. BMC Bioinformatics, 2011, 12, I1.	2.6	0
35	PDC - a probabilistic distributional clustering algorithm: a case study on suicide articles in PubMed. AMIA Summits on Translational Science Proceedings, 2020, 2020, 259-268.	0.4	0